④

SPHERICALLY INVARIANT PROCESSES:

THEIR NONLINEAR STRUCTURE, DISCRIMINATION, AND ESTIMATION*

by

Steel T. Huang

Department of Mathematics
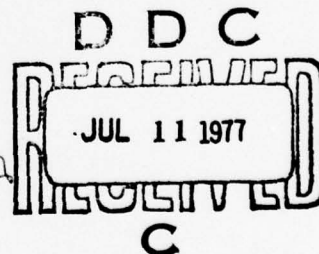University of Cincinnati
Cincinnati, Ohio 45221

and

Stamatis Cambanis

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina 27514

ABSTRACT

The structure of the nonlinear space of a spherically invariant

process is studied and the problem of discriminating between two

spherically invariant processes as well as the problem of nonlinear

estimation in spherically invariant processes are solved.

D D C
RECEIVED
JUL 11 1977
C

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) AFOSR-TR-77-0708	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SPHERICALLY INVARIANT PROCESSES: THEIR NONLINEAR STRUCTURE, DISCRIMINATION AND ESTIMATION,		5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Steel T. Huang Stamatis Cambanis (15)		8. CONTRACT OR GRANT NUMBER(s) AF- AFOSR 75-2796 - 75
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of North Carolina Department of Statistics Chapel Hill, NC 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE 1977
		13. NUMBER OF PAGES 39
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

spherically invariant processes; Gaussian processes; nonlinear space; discrimination between two spherically invariant processes; nonlinear estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The structure of the nonlinear space of a spherically invariant process is studied and the problem of discriminating between two spherically invariant processes as well as the problem of nonlinear estimation in spherically invariant processes are solved.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

0.   INTRODUCTION

This paper attempts a systematic study of spherically invariant pro-
cesses (SIP's), i.e., processes whose finite dimensional distributions are
mixtures of Gaussian distributions. Section 1 contains the basic properties
of SIP's, including their representation in terms of Gaussian processes,
which are used throughout. The structure of the nonlinear space of a second
order SIP is considered in Section 2. Section 3 solves the problem of
discriminating between two second order SIP's, and Section 4 the nonlinear
estimation problem for second order SIP's and in particular for Gaussian
processes.

Our basic notation and terminology is as follows. $X = (X_t, t \epsilon T)$ is
a stochastic process defined on a probability space $(\Omega, \mathcal{B}, P)$. T is an arbitrary
index set; sometimes it is taken to be a real line interval, but since this is
clear from the context it is not emphasized. $\mathcal{B}$ is usually taken to be $\mathcal{B}(X)$,
the σ-field generated by the random variables of the process X, or $\overline{\mathcal{B}}(X)$, the
completion of $\mathcal{B}(X)$ with respect to P. For each $\omega \epsilon \Omega$, $X(\omega)$ is the corresponding
sample path of the process X, which is an element of $\mathbb{R}^T$, the space of all
functions defined on T. X induces a probability measure $\mu = P \cdot X^{-1}$ on
$(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$, where $\mathcal{B}(\mathbb{R}^T)$ is the σ-field generated by the cylinder sets of $\mathbb{R}^T$.
X is a coordinate process if $(\Omega, \mathcal{B}, P) = (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \mu)$ and $X_t(\omega) = \omega(t)$. The
nonlinear space of X, $L_2(X) = L_2(\Omega, \mathcal{B}(X), P)$, is the set of all $\mathcal{B}(X)$-measurable
random variables with finite second moment which are called (nonlinear) L_2-
functionals of X. X is a second order process if $EX_t^2 < \infty$ for all $t \epsilon T$. The
linear space of a second order process X, H(X), is the closed subspace of
$L_2(X)$ spanned by X_t, $t \epsilon T$, and its elements are called linear L_2-functionals of
X.

1. SPHERICALLY INVARIANT PROCESSES (SIP's)

It is well known that all mean square estimation problems on Gaussian processes have linear solutions and that Gaussian processes are closed under linear operations. Vershik (1964) showed that these two properties do not uniquely characterize the Gaussian processes. They do, however, characterize the class of SIP's.

Let $X = (X_t,\ t \epsilon T)$ be a second order process with mean $m(t)$ and covariance function $r(t,s)$. Then X is said to be an SIP if all r.v.'s in $H(X-m)$ having the same variance have the same distribution. Also, X is called degenerate if $H(X-m)$ is finite dimensional. Since the nonlinear structure of a non-degenerate and a degenerate SIP are considerably different, we will restrict our present investigation to the nondegenerate case; the finite dimensional case will be treated elsewhere. A SIP X is a mixture of Gaussian processes. It can be determined by its mean $m(t)$, a covariance function $R(t,s)$, and a probability distribution $F(\alpha)$ on $\mathbb{R}^+$; the characteristic function of $X_{t_1}, \ldots, X_{t_k}$ $(t_1, \ldots, t_k \epsilon T)$ is given by

$$(1.1) \qquad \int e^{-\frac{\alpha}{2} \Sigma R(t_i, t_j)(u_i - m(t_i))(u_j - m(t_j))} dF(\alpha)$$

(Vershik (1964) and Nagornyi (1970)). In order to avoid the trivial case where X is a constant process, we will assume throughout that $F(0^+) < 1$. Such an SIP determined by m, R and F will be denoted, in short, by SIP(m,R;F). A probability measure μ on the sample space $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ is said to be a *spherically invariant measure* (SIM) if it is induced by a SIP, or, equivalently, if the coordinate process on $(\mathbb{R}^t, \mathcal{B}(\mathbb{R}^T), \mu)$ is a SIP. A SIM induced by a SIP (m, R, F) will be denoted by SIM(m,R;F). When F puts all its mass at the point 1 the SIP (m,R;F) and the SIM (m,R;F) become

Gaussian with mean m and covariance R; we will use the notation GP(m,R) and GM(m,R) respectively.

LEMMA 1.1   Let μ be an SIM (m,R:F) and for each $\alpha \geq 0$ let μ_α be a GM(m,αR).   *Then*

(1.2) $$\mu(E) = \int \mu_\alpha(E) dF(\alpha) \qquad \text{for all } E \epsilon \mathcal{B}(\mathbb{R}^T).$$

Furthermore, for any measurable function θ on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \mu)$ which is non-negative or integrable we have

(1.3) $$E\theta = \int E_\alpha \theta dF(\alpha)$$

where $E\theta = \int \theta d\mu$ and $E_\alpha \theta = \int \theta d\mu_\alpha$.

<u>Proof:</u> For each fixed cylinder set E, $\mu_\alpha(E)$ is a measurable function of α; and the family of sets E such that $\mu_\alpha(E)$ is α-measurable is a σ-field. There-fore, $\mu_\alpha(E)$ is α-measurable for every $E\epsilon\mathcal{B}(\mathbb{R}^T)$. Thus $\int \mu_\alpha(E)dF(\alpha)$ is well-defined and is easily checked to be a probability measure. The characteristic functions of its finite dimensional distributions are given by (1.1), and since they uniquely determine a probability measure on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$, we have (1.2).

(1.3) holds for $\theta = 1_E$ $(E\epsilon\mathcal{B}(\mathbb{R}^T))$ by (1.2). Hence it holds for simple functions θ, and thus for θ nonnegative by considering a sequence of simple functions increasing to θ. For θ integrable, consider the positive and nega-tive parts separately. $\square$

It follows from Lemma 1.1 that if X is a coordinate SIP (m,R;F) then, under each μ_α, X is a GP(m,αR). The following theorem generalizes this fact.

THEOREM 1.2. Let X be a SIP (m,R;F) on $(\Omega, \mathcal{B}(X), P)$. Then for each $\alpha \geq 0$ there exists a probability measure P_α on $(\Omega, \mathcal{B}(X))$ such that $\mu_\alpha = P_\alpha \cdot X^{-1}$ is a GM(m,αR). Hence under each P_α, X is a GP(m,αR) and the corresponding

formulae (1.2) and (1.3) hold.

Proof: Let μ be the SIM induced on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ by X and let μ_α, $\alpha \geq 0$, be GM(m, R) as in Lemma 1.1. Let $C(\mathbb{R}^T)$ be the field of all cylinder sets in $\mathcal{B}(\mathbb{R}^T)$ and $C(X) = X^{-1} \cdot C(\mathbb{R}^T) \subset \mathcal{B}(X)$. For each $D \in C(X)$, define $P_\alpha(D) = \mu_\alpha(E)$ if $D = X^{-1}(E)$. We claim that P_α is well defined. It is not hard to see (by looking at their characteristic functions) that when restricted to a finite dimensional space μ_α is absolutely continuous with respect to μ. Thus if $D = X^{-1}(E) = X^{-1}(E')$, $E, E' \in C(\mathbb{R}^T)$, then $\mu(E) = P(D) = \mu(E_1)$ and hence $\mu_\alpha(E) = \mu_\alpha(E')$. Consequently, P_α is a probability measure on the field $C(X)$ and hence it has a unique extension to a probability measure on $\mathcal{B}(X)$. Now it follows that $\mu_\alpha = P_\alpha \cdot X^{-1}$ since the two measures coincide on $C(\mathbb{R}^T)$. □

(1.2) in Theorem 1.2 was proved by Gualtierotti (1974) for P an SIM on a separable Hilbert space. Note that since X is of second order, it follows from

$$\infty > EX_t^2 = \int E_\alpha X_t^2 dF(\alpha) = \int [m^2(t) + \alpha R(t,t)] dF(\alpha)$$

that

$$\alpha_1 = \int \alpha dF(\alpha) < \infty .$$

THEOREM 1.3. Let X be a SIP $(0,R;F)$ and $\{\xi_n\}$ a sequence in H(X). Then there exist versions of the ξ_n's such that under each P_α the r.v.'s $\{\xi_n\}$ are jointly Gaussian with zero mean and covariance

(1.4) $$E_\alpha \xi_n \xi_m = \frac{\alpha}{\alpha_1} E\xi_n \xi_m = \alpha E_1 \xi_n \xi_m \ ,$$

and, moreover, for any measurable function g on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ we have

(1.5) $$E_\alpha g(\xi_1, \xi_2, \ldots) = E_1 g(\sqrt{\alpha}\ \xi_1, \ \sqrt{\alpha}\ \xi_2, \ldots)$$

whenever one of these expectations exists.

<u>Proof:</u>   In view of Theorem 1.2 and the fact that every measurable function $\theta(\omega)$ on $(\Omega, \mathcal{B}(X))$ is of the form $f(X(\omega))$ for some measurable function f on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$, it suffices to prove the assertion for X a coordinate process. For each r.v. $\xi_n(x)$, $x \in \mathbb{R}^T$, there is a sequence $\ell_i^{(n)}(x)$ of finite linear combinations of $\{f_t(x) = x(t), t \in T\}$ such that $\xi_n(x) = \lim_i \ell_i^{(n)}(x)$ a.s. $[\mu]$. Let $C_n = \{x \in \mathbf{R}^T : \ell_i^{(n)}(x) \to \cdot$ as $i \to \infty\}$ and let $C = \cap C_n$. C is clearly a measurable linear space having μ probability 1. Now take each ξ_n to be $\lim_i \ell_i^{(n)}(x)$ for $x \in C$ and 0 for $x \notin C$. Since $1 = \mu(C) = \int \mu_\alpha(C) dF(\alpha)$ by Lemma 1.1, and $F(0^+) < 1$, there is $\alpha_0 > 0$ such that $\mu_{\alpha_0}(C) = 1$. Let Y be a $G(0, R)$ on some probability space $(\Omega_0, \mathcal{B}_0, Q)$. Then since C is a linear space we have for all $\alpha \geq 0$

$$\mu_\alpha(C) = Q(\sqrt{\alpha} \ Y \epsilon C) = Q(\sqrt{\alpha_0} \ Y \epsilon C) = \mu_{\alpha_0}(C) = 1 \ .$$

Thus $\xi_n = \lim_i \ell_i^{(n)}$ a.e. $[\mu_\alpha]$, for all $\alpha \geq 0$. Since under each μ_α, $\{\ell_i^{(n)}, n \geq 1, i \geq 1\}$ is a Gaussian family, it follows that $\{\xi_n\}$ are jointly Gaussian under each μ_α.

To show (1.4) we recall that for a Gaussian family a.s. convergence is equivalent to mean square convergence. Thus

$$E_\alpha \xi_n \xi_m = \lim_{i,j} E_\alpha \ell_i^{(n)} \ell_j^{(m)}$$
$$= \lim_{i,j} \alpha E_1 \ell_i^{(n)} \ell_j^{(m)} = \alpha E_1 \xi_i \xi_j \ ,$$

and hence $E \xi_i \xi_j = \int E_\alpha \xi_i \xi_j dF(\alpha) = \alpha_1 E_1 \xi_i \xi_j = \frac{\alpha_1}{\alpha} E_\alpha \xi_i \xi_j$.

It is easy to see that $\{\xi_n\}$ under μ_α has the same probability law as $\{\sqrt{\alpha} \ \xi_n\}$ under μ_1. (1.5) is now evident.   □

We now introduce a r.v. A which will play a central role in the representation of a SIP as well as in the study of its nonlinear space.

Pick an orthogonal sequence $\{\xi_n\}$ from $H(X)$ with $E\xi_n^2 = \alpha_1$. Then by Theorem 1.3 we may assume that under each P_α, $\{\xi_n\}$ is a sequence of independent zero mean Gaussian r.v.'s with $E_\alpha \xi_n^2 = \alpha$. Define

$$A_n = \frac{1}{n} \sum_1^n \xi_i^2 \quad .$$

By the law of large numbers, we have $\lim A_n = \alpha$ a.s. $[P_\alpha]$. Let $C^* = \{\omega \in \Omega : A_n(\omega) \to \cdot\}$. Then

$$P(C^*) = \int P_\alpha(C^*) dF(\alpha) = \int dF(\alpha) = 1 \quad .$$

Thus A_n converges a.s. $[P]$. Let

$$(1.6) \qquad A(\omega) = \begin{cases} \lim A_n(\omega) & \text{if } \omega \in C^* \\ 0 & \text{if } \omega \notin C^* \quad . \end{cases}$$

Then

$$(1.7) \qquad A = \alpha \quad \text{a.s. } [P_\alpha] \quad ,$$

and the distribution function of A is F since

$$(1.8) \qquad P(A \le a) = \int P_\alpha(A \le a) dF(\alpha) = \int 1_{[0,a]}(\alpha) dF(\alpha) = F(a) .$$

If we put $\Omega_\alpha = \{\omega \in \Omega : A(\omega) = \alpha\}$ then $P_\alpha(\Omega_\alpha) = 1$, and thus the probability measures P_α, $\alpha \ge 0$, are mutually singular (which is of course well known).

We now arrive at the main theorem of this section.

THEOREM 1.4. *A nondegenerate second order process* X *on* $(\Omega, \mathcal{B}(X), P)$ *is a SIP* $(0;R;F)$ *if and only if it has the representation*

$$(1.9) \qquad X_t = A^{1/2} Y_t \quad \text{a.s. for all } t \in T$$

where A *is a nonnegative r.v. with distribution function* F *and, conditioned on* $A > 0$, Y *is a nondegenerate* $GP(0,R)$ *independent of* A.

Proof: The "if" part is clear. To show the "only if" part, let A be the r.v. defined by (1.6) and let

$$(1.10) \qquad Y = \begin{cases} \dfrac{X}{A^{1/2}} & \text{if } A > 0 \\ \\ 0 & \text{if } A = 0 . \end{cases}$$

Then it follows from Theorem 1.2 and (1.7) that $P(X_t = A^{1/2}Y_t) = \int P_\alpha(X_t = \alpha^{1/2}Y_t)dF(\alpha) = 1$; and that Y is a GP(0,R) under each P_α for $\alpha > 0$ and thus also under $\overline{P}(\cdot) = (1-F(0^+))^{-1}\int_{(0,\infty)} P_\alpha(\cdot)dF(\alpha)$, which is just the conditional probability of P given $A > 0$. Now for $D \in \mathcal{B}(\mathbb{R})$ and $E \in \mathcal{B}(\mathbb{R}^T)$ we have

$$\overline{P}(A \in D,\ Y \in E) = \frac{1}{1-F(0^+)} \int_{(0,\infty)} 1_D(\alpha)P_\alpha(Y \in E)dF(\alpha)$$

$$= \frac{1}{1-F(0^+)} P(A \in D) \cdot \overline{P}(Y \in E)$$

$$= \overline{P}(A \in D) \quad \overline{P}(Y \in E)$$

since $P_\alpha(Y \in E) = \overline{P}(Y \in E)$ for $\alpha > 0$. Hence A and Y are conditionally independent given $A > 0$. □

The representation (1.9) of a SIP was first noted by Besson (1974) who, without constructing the r.v. A, showed its existence by employing a result in Bretagnolle et. al (1966) concerning symmetrically dependent (exchangeable) r.v.'s. Our approach seems more elementary and direct, and also yields further results.

(1.9) reveals that a SIP is merely a conditional Gaussian process. (More specifically, given $A = \alpha$, the SIP (0,R;F) is a GP (0,αR).) By requiring $F(0^+) = 0$, Theorem 1.4 may be stated in a more appealing way.

THEOREM 1.5. *A nondegenerate second order process* X *on* $(\Omega, \mathcal{B}(X), P)$ *is a* *SIP* $(0, R; F)$ *satisfying* $F(0^+) = 0$ *if and only if it has the representation*

$$X_t = A^{1/2} Y_t \quad a.s., \ t \in T$$

where Y *is a nondegenerate GP* $(0, R)$ *and* A *is a positive r.v. independent of* Y *having distribution function* F.

Theorem 1.4 enables one to read off many properties of SIP's immediately. For instance a SIP $(0, R; F)$ X is of order $p > 0$, i.e., $E|X_t|^p < \infty$, if and only if $EA^{p/2} = \int \alpha^{p/2} dF(\alpha) < \infty$. If X is of order 2, then continuity in probability of X is equivalent to mean square continuity [Besson (1974)]. All the usual (local and global) analytic properties of the sample functions of X depend only on R and not on F, while properties of maxima and crossings depend on both R and F. Kallianpur's zero-one law and Slepian's lemma take the following form for SIP's. (The proofs are straightforward and are thus omitted.)

COROLLARY 1.6. *If* μ *is an SIM* $(0, R; F)$ *and* L *a* $\mathcal{B}(\mathbb{R}^T)$-*measurable linear subspace of* $\mathbb{R}^T$, *then* $\mu(L) = F(0^+)$ *or* 1. *Furthermore,* $\mu(L) = 1$ *if and only if for some* $\alpha > 0$, L *is* $\overline{\mathcal{B}}_\alpha(\mathbb{R}^T)$-*measurable and* $\mu_\alpha(L) = 1$.

In Corollary 1.6 μ_α are as in Lemma 1.1 and $\overline{\mathcal{B}}(\mathbb{R}^T)$, resp. $\overline{\mathcal{B}}_\alpha(\mathbb{R}^T)$, denotes the completion of $\mathcal{B}(\mathbb{R}^T)$ with respect to μ, resp. μ_α.

COROLLARY 1.7. *If* X_i *is a separable SIP* $(0, R_i; F_i)$, $i = 1, 2$, *and if*

$$R_1(t,t) = R_2(t,t) \quad , \quad R_1(t,s) \leq R_2(t,s) \quad \text{for all } t, s \in T ,$$
$$F_1(0^+) = F_2(0^+) \quad , \quad F_1(\alpha) \leq F_2(\alpha) \quad \text{for all } \alpha > 0 ,$$

then for all u,

$$P\{\sup_{t \in T} X_1(t) < u\} \leq P\{\sup_{t \in T} X_2(t) < u\}.$$

We now give an interesting example of a sample continuous martingale whose family of σ-fields is not continuous. Let $W = \{W_t, 0 \leq t < \infty\}$ be a Wiener process, and let $X = A^{1/2}W$ where A is a nonnegative r.v. independent of W and whose σ-field is nontrivial. It is easily checked that X is a sample continuous martingale. Let $\mathcal{B}_t = \mathcal{B}(X_s, 0 \leq s \leq t)$, $0 \leq t < \infty$. We will show that $\mathcal{B}_t$ is not continuous at $t = 0$. Fix $t > 0$ and consider the quadratic variation $M(t)$ of X over the interval $[0,t]$. We have

$$M(t) = \lim_{n} \sum_{j=1}^{2^{n}-1} [X(\frac{j+1}{2^n} t) - X(\frac{j}{2^n} t)]^2$$

$$= A \lim_{n} \sum_{j=1}^{2^{n}-1} [W(\frac{j+1}{2^n} t) - W(\frac{j}{2^n} t)]^2$$

$$= At \quad a.s.$$

where the last equality is a theorem of Lévy [Doob (1953)]. This implies that A is $\mathcal{B}_t$-measurable for all $t > 0$.

2. THE NONLINEAR SPACE OF A SIP.

In this section we study the structure of the nonlinear space $L_2(X)$ of a SIP$(O,R;F)$ X, using the canonical representation of X in Theorem 1.4 and the well known properties of the nonlinear space of a GP. When X is second order, the relation between the linear space H(X) and the nonlinear space $L_2(X)$ is shown in Theorem 2.4 and complete orthonormal sets (CONS's) in $L_2(X)$ are given in Theorems 2.2 and 2.3. When F has a moment generating function, Theorem 2.5 shows that $L_2(X)$ has the orthogonal decomposition $\oplus_{p \geq 0} H_p(X)$, where $H_p(X)$ is the p-th homogeneous chaos of X; and Theorem 2.6 shows the relation between $H_p(X)$ and H(X) and gives CONS's in each $H_p(X)$.

Theorems 2.2 to 2.4 are based on the following property.

LEMMA 2.1 Let B_1 and B_2 be two independent σ-fields on a probability space (Ω, B, P) such that B is generated by B_1 and B_2. Then

$$L_2(\Omega, B, P) \cong L_2(\Omega, B_1, P) \otimes L_2(\Omega, B_2, P)$$

under the correspondence $fg \leftrightarrow f \otimes g$.

Proof: We will write $L_2(B_i)$ for $L_2(\Omega, B_i, P)$. Consider the mapping taking $f \otimes g$ to fg for all $f \in L_2(B_1)$, $g \in L_2(B_2)$. From the independence of B_1 and B_2 it follows that $fg \in L_2(B)$ and

$$< f_1 \otimes g_1, f_2 \otimes g_2 >_{L_2(B_1) \otimes L_2(B_2)} = < f_1 g_1, f_2 g_2 >_{L_2(B)}$$

Thus the mapping preserves inner products and it has a unique extension to an isomorphism between the closed subspace spanned by $\{f \otimes g\}$, which is in fact $L_2(B_1) \otimes L_2(B_2)$, and the closed subspace of $L_2(B)$ spanned by $\{fg\}$. The assertion is proved if we show that $M = \{fg : f \in L_2(B_1), g \in L_2(B_2)\}$ is a complete set in $L_2(B)$.

Consider the family $\mathcal{B}^*$ of sets of the form $\cup_1^n (E_i \cap F_i)$, $E_i \in \mathcal{B}_1$ and $F_i \in \mathcal{B}_2$. Since $\mathcal{B}_1$ and $\mathcal{B}_2$ are fields and they generate $\mathcal{B}$, it is readily seen that $\mathcal{B}^*$ is a field generating $\mathcal{B}$. We may even assume, after a moment's reflection, that the sets $E_i \cap F_i$, $1 \le i \le n$, are disjoint. Now a standard monotone class argument shows that the set $\{1_{\cup_1^n (E_i \cap F_i)} = \sum_1^n 1_{E_i} 1_{F_i}\}$

is complete in $L_2(\mathcal{B})$. Thus M is complete. □

Let $X = \{X_t, t \in T\}$ be a second order nondegenerate $SIP(0,R;F)$ and $X = A^{\frac{1}{2}} Y$ its canonical representation of Theorem 1.4. Let $\{\xi_\gamma, \gamma \in \Gamma\}$ (Γ linearly ordered) and $\{e_n, 1 \le n \le N\}$ (N may be infinite) be CONS's in $H(X)$ and $L_2(dF)$ respectively.

THEOREM 2.2 If $F(0^+) = 0$ *then the family*

(2.1) $\quad e_n(A) \cdot (\frac{1}{p_1! \cdots p_k!})^{\frac{1}{2}} H_{p_1}(((\frac{\alpha_1}{A})^{\frac{1}{2}} \xi_{\gamma_1}) \ldots H_{p_k}((\frac{\alpha_1}{A})^{\frac{1}{2}} \xi_{\gamma_k})$

$\quad = e_n(A) \cdot (\frac{1}{p_1! \cdots p_k!})^{\frac{1}{2}} (\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p_1, \frac{A}{\alpha_1}}(\xi_{\gamma_1}) \ldots H_{p_k, \frac{A}{\alpha_1}}(\xi_{\gamma_k})$

where $1 \le n \le N$, $k \ge 1$, $p_1 + \ldots + p_k = p \ge 0$, $\gamma_1 < \ldots < \gamma_k$ *in* Γ, *is a CONS in* $L_2(X)$

Proof: Assume $F(0^+) = 0$. Then $\mathcal{B}(A)$ and $\mathcal{B}(Y)$ are independent σ-fields generating $\mathcal{B}(X)$. By Lemma 2.1 we have

(2.2) $\qquad L_2(X) \tilde{=} L_2(A) \otimes L_2(Y)$.

It is easily verified that $A^{\frac{1}{2}} H(Y) = H(X)$ and $\{\eta_\gamma = (\frac{\alpha_1}{A})^{\frac{1}{2}} \xi_\gamma, \gamma \in \Gamma\}$ is complete in $H(Y)$. (Recall that Y is a GP.) Indeed $\{\eta_\gamma, \gamma \in \Gamma\}$ is a CONS in $H(Y)$ since

$$E(\frac{\alpha_1}{A} \xi_\gamma \xi_\beta) = \int \frac{\alpha_1}{\alpha} E_\alpha \xi_\gamma \xi_\beta dF(\alpha) = E \xi_\gamma \xi_\beta$$

(by Theorem 1.3). Now (2.2) and the celebrated theorem of Cameron and Martin yield the result. □

We remark that in dealing with a SIP, the case $F(0+) = 0$
(i.e. X cannot be zero with positive probability) is of main interest.

We now show how to obtain the counterpart of Theorem 2.2 when
$0 < F(0+) < 1$. Let $\Omega_1 = \{\omega \in \Omega : A(\omega) = 0\}$, $\Omega_2 = \{\omega \in \Omega : A(\omega) > 0\}$, and
consider the restrictions of $(\Omega, B(X), P)$ to Ω_1 and Ω_2 respectively,
$(\Omega_1, B_{\Omega_1}(X), P_{\Omega_1})$ and $(\Omega_2, B_{\Omega_2}(X), P_{\Omega_2})$. Then

$$L_2(X) = L_2(\Omega_1, B_{\Omega_1}(X), P_{\Omega_1}) \oplus L_2(\Omega_2, B_{\Omega_2}(X), P_{\Omega_2}) .$$

It is easily seen that $\qquad L_2(\Omega_1, B_{\Omega_1}(X), P_{\Omega_1}) \cong \mathbb{R}$,
where $\mathbb{R}$ is equiped with the usual inner product ($<x, y>_{\mathbb{R}} = xy$). Also $P_{\Omega_2} = Q_{\Omega_2}$ where Q is defined on $B(X)$ by $Q(\circ) = \int_{(0,\infty)} P_\alpha(\circ) dF(\alpha)$, and thus

$$L_2(\Omega_2, B_{\Omega_2}(X), P_{\Omega_2}) \cong L_2(\Omega, B(X), \overline{P} = \frac{Q}{1-F(0+)})$$

under the correspondence $(1-F(0+))^{-\frac{1}{2}} 1_{\Omega_2} f \leftrightarrow f$. But under $\overline{P}$, X is a
SIP$(0, R; G)$ where $G(\alpha) = (1-F(0+))^{-\frac{1}{2}}(F(\alpha)-F(0+))$, and Theorem 2.2 yields
the following CONS.

$\underline{THEOREM\ 2.3}$ If $0 < F(0+) < 1$ and if $\{e_n, 1 \leq n \leq N\}$ is a CONS in $L_2(d\{F-F(0+)\})$
then the family

(2.3)

$$(\frac{1}{F(0+)})^{\frac{1}{2}} 1_{\{0\}}(A),$$

$$(\frac{1}{1-F(0+)})^{\frac{1}{2}} 1_{(0,\infty)}(A)\ e_n(A)\ (\frac{1}{p_1!\ldots p_k!})^{\frac{1}{2}}(\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p_1, \frac{A}{\alpha_1}}(\xi_{\gamma_1})\ldots H_{p_k, \frac{A}{\alpha_1}}(\xi_{\gamma_k})$$

where $1 \leq n \leq N$, $k \geq 1$, $p_1 + \ldots + p_k = p \geq 0$, $\gamma_1 < \ldots < \gamma_k$ in Γ, is a CONS in
$L_2(X)$.

Consider the correspondence

$$(2.4) \quad e_n(A) \; (\frac{1}{p_1!\ldots p_k!})^{\frac{1}{2}} (\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p_1,\frac{A}{\alpha_1}}(\xi_{\gamma_1}) \ldots H_{p_k,\frac{A}{\alpha_1}}(\xi_{\gamma_k}) \leftrightarrow$$

$$e_n(A) \; (\frac{p!}{p_1!\ldots p_k!})^{\frac{1}{2}} (\xi_{\gamma_1}^{\tilde{\otimes}\rho_1} \tilde{\otimes}\ldots\tilde{\otimes} \xi_{\gamma_k}^{\tilde{\otimes}\rho_k})$$

when $F(0+) = 0$, and when $0 < F(0+) < 1$ the correspondence

$$1 \leftrightarrow (\frac{1}{F(0+)})^{\frac{1}{2}} \; 1_{\{0\}}(A),$$

$$(2.5)$$

$$(\frac{1}{1-F(0+)})^{\frac{1}{2}} 1_{(0,\infty)}(A) \cdot e_n(A) (\frac{1}{p_1!\ldots p_k!})^{\frac{1}{2}}(\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p_1,\frac{A}{\alpha_1}}(\xi_{\gamma_1})\ldots H_{p_k,\frac{A}{\alpha_1}}(\xi_{\gamma_k})$$

$$\leftrightarrow e_n \otimes (\frac{p!}{p_1!\ldots p_k!})^{\frac{1}{2}} (\xi_{\gamma_1}^{\tilde{\otimes}p_1} \tilde{\otimes}\ldots\tilde{\otimes} \xi_{\gamma_k}^{\tilde{\otimes}p_k}).$$

Since $(\frac{p!}{p_1!\ldots p_k!})^{\frac{1}{2}} (\xi_{\gamma_1}^{\tilde{\otimes}p_1} \tilde{\otimes}\ldots\tilde{\otimes} \xi_{\gamma_k}^{\tilde{\otimes}\rho_k})$, where $k \geq 1$, $p_1 +\ldots+ p_k = p \geq 0$,

$\gamma_1 <\ldots< \gamma_k$, form a CONS in $H^{\tilde{\otimes}p}(X)$, we have

THEOREM 2.4

$$L_2(X) \overset{\sim}{=} L_2(dF) \otimes (\underset{p\geq 0}{\oplus} H^{\tilde{\otimes}p}(X))$$

$$\overset{\sim}{=} L_2(A) \otimes (\underset{p\geq 0}{\oplus} H^{\tilde{\otimes}p}(X)), \; if \; F(0+) = 0 \; ;$$

$$L_2(X) \overset{\sim}{=} \mathbb{R} \oplus \{L_2(d(F-F(C+)) \otimes (\underset{p\geq 0}{\oplus} H^{\tilde{\otimes}p}(X))\}, if \; 0 < F(0^+) < 1.$$

Suppose that every $\xi \in H(X)$ has all moments finite. Let $P(X)$ be the
linear space of all polynomials in elements of $H(X)$ and let $P_p(X)$ $(p\geq 0)$
be the linear space of all polynomials in $P(X)$ of degree at most p; hence
$P_0(X)$ is the set of all constants. Let $Q_0(X) = P_0(X)$ and for $p\geq 1$ let
$Q_p(X)$ be the set of all polynomials in $P_p(X)$ orthogonal to $P_{p-1}(X)$. Denote

by $H_p(X)$ the closure of $Q_p(X)$ in $L_2(X)$. $Q_p(X)$ is called the p-th polynomial chaos and $H_p(X)$ is called the p-th homogenous chaos.

When X is a SIP(O,R;F), in order to have all moments of $\xi \epsilon H(X)$ finite we introduce the following "moment" condition:

(M) The moment generating function of F exists, i.e.

$$\int e^{\alpha t} \, dF(\alpha) < \infty \quad \text{for all } t \epsilon \mathbb{R}.$$

Under the condition (M) we have for $\xi \epsilon H(X)$ (by (1.5)),

$$E|\xi|^p = \int E_\alpha |\xi|^p \, dF(\alpha) = \int \alpha^{p/2} E_1 |\xi|^p dF(\alpha) = \alpha_{p/2} E_1 |\xi|^p < \infty$$

where $\alpha_p = \int \alpha^p dF(\alpha)$, $p \geq 0$.

THEOREM 2.5 If (M) holds,

$$L_2(X) = \bigoplus_{p \geq 0} H_p(X).$$

Proof: It is well-known (see Neveu (1968)) that if $e^{|\xi|} \epsilon L_2(X)$ for every $\xi \epsilon H(X)$ then $L_2(X) = \bigoplus_{p \geq 0} H_p(X)$ (for arbitrary process X). Thus it suffices to show that $e^{|\xi|}$ is integrable if (M) holds. Since under each P_α, ξ is a zero mean Gaussian variable with variance $\sigma^2 = \frac{\alpha}{\alpha_1} E\xi^2$, we have

$$E_\alpha e^{|\xi|} = \frac{2}{\sqrt{2\pi} \, \sigma} \int_0^\infty e^{x - (x^2/2\sigma^2)} \, dx$$

$$= \frac{2e^{\sigma^2/2}}{\sqrt{2\pi} \, \sigma} \int_{-\sigma^2}^\infty e^{-x^2/2\sigma^2} \, dx \leq 2e^{\sigma^2/2}.$$

Thus, under (M),

$$Ee^{|\xi|} = \int E_\alpha e^{|\xi|} \, dF(\alpha) \leq 2 \int e^{\frac{\alpha}{2\alpha_1} E\xi^2} \, dF(\alpha) < \infty \, . \qquad \square$$

We now establish the relationship between the decomposition of Theorem 2.5 and the representation of Theorem 2.4. Assume the condition (M). Then $L_2(X) = \underset{p \geq 0}{\oplus} H_p(X)$, and $\{A^n, 0 \leq n < \infty\}$ is a complete set of $L_2(A)$. For each Fixed $q \geq 0$, applying the Gram-Schmidt onthonormalization procedure, to $\{A^n\}$ with respect to the inner product $<A^m, A^n>_q = E \, A^{m+n+q}$ we obtain the set $\{e_n^q(A), 0 \leq n < \infty\}$ which is complete in $L_2(A)$ and orthonormal relative to $<\cdot,\cdot>_q$. Note that $e_n^q(A)$ is a polynomial in A with degree n.

Now assume $F(0+) = 0$. By Theorem 2.4 we have $L_2(A) \otimes (\underset{p \geq 0}{\oplus} H^{\widetilde{\otimes p}}(X)) \cong L_2(X)$. Denote this isomorphism by Φ. Let

$$(2.6) \qquad H_{n,q}(X) = \Phi\{e_n^q(A) \left(\frac{A}{\alpha_1}\right)^{q/2} \otimes H^{\widetilde{\otimes q}}(X)\} \, .$$

It is not difficult to see from (2.4) that

$$(2.7) \qquad H_{n,q}(X) \subset \overline{P}_{2n+q}(X) \setminus \overline{P}_{2n+q-1}(X)$$

since $A \in \overline{P}_2$, from the definition of A.

THEOREM 2.6 Under conditions (M) and $F(0+) = 0$,

$$(2.8) \qquad H_p(X) = \underset{2n+q=p}{\oplus} H_{n,q}(X)$$

and a CONS in $H_p(X)$ is given by

$$(2.9) \qquad \left(\frac{\alpha_1^q}{q_1! \ldots q_k!}\right)^{\frac{1}{2}} e_n^q(A) \, H_{q_1,\frac{A}{\alpha_1}}(\xi_{\gamma_1}) \ldots H_{q_k,\frac{A}{\alpha_1}}(\xi_{\gamma_k})$$

where $2n+q = p$, $k \geq 1$, $q_1 + \ldots + q_k = q$, $\gamma_1 < \ldots < \gamma_k$.

<u>Proof</u>: First we show that $H_{n,q}(X) \perp H_{n',q'}(X)$ if $(n,q) \neq (n',q')$. It is clear from (2.6) that $H_{n,q}(X) \perp H_{\underset{\sim}{n}',q'}(X)$ if $q \neq q'$. Suppose now that $q = q'$ and $n \neq n'$. For $\theta, \theta' \in H^{\widetilde{\otimes}q}(X)$ we have

$$E\{ \Phi(e_n^q(A) (\tfrac{A}{\alpha_1})^{q/2} \otimes \theta) \quad \Phi(e_{n'}^q(A) (\tfrac{A}{\alpha_1})^{q/2} \otimes \theta') \}$$

$$= \frac{1}{\alpha_1^q} E\{e_n^q(A) e_{n'}^q(A) A^q\} <\theta, \theta'>_{H^{\widetilde{\otimes}q}(X)} = 0$$

since $<e_n^q, e_{n'}^q>_q = 0$ and thus $H_{n,q}(X) \perp H_{n'q}(X)$.

In order to show (2.8) it suffices to show that $\overline{P}_p(X) = \underset{2n+q\leq p}{\oplus} H_{n,q}(X)$ because of (2.7). We need to show that $\xi_1^{r_1}\dots\xi_K^{r_k} \in \underset{2n+q\leq p}{\oplus} H_{n,q}(X)$ for all $\xi_1,\dots,\xi_k \in H(X)$ and $r_1+\dots+r_k \leq p$. For ease of exposition we show this for ξ^r only. Write $\xi = A^{\frac{1}{2}}\eta, \eta \in H(Y)$. Then

$$\xi^r = A^{r/2}\eta^r = A^{r/2}\{H_{r,E\eta^2}(\eta) + \text{Const.} \, H_{r-2,E\eta^2}(\eta) + \text{Const.} \, H_{r-4,E\eta^2}(\eta)+\dots\}$$

$$= H_{r,AE\eta^2}(\xi) + \text{Const.} \, A \, H_{r-2,AE\eta^2}(\xi) + \text{Const.} \, A^2 H_{r-4,AE\eta^2}(\xi)+\dots \, .$$

Note that $A^m H_{r-2m,AE\eta^2}(\eta) \in \underset{n=0}{\overset{m}{\oplus}} H_{n,r-2m}(X) \subset \underset{2n+q\leq r}{\oplus} H_{n,q}(X)$. Thus $\xi^r \in \underset{2n+q\leq p}{\oplus} H_{n,p}(X)$ and (2.8) is proved.

(2.9) follows from (2.4), (2.6), (2.8) and $||e_n^q(A) (\tfrac{A}{\alpha_1})^{q/2}||^2 = \alpha_1^{-q}$. $\square$

Of course $H_1(X) = H(X)$ and, for instance $H_2(X) = H_{1,0}(X) \oplus H_{0,2}(X)$ where $H_{1,0}(X) = \overline{sp}\{A-\alpha_1\}$ and

$$H_{0,2}(X) = \overline{sp}\{ (\frac{\alpha_1^2}{p_1!p_2!})^{\frac{1}{2}} H_{p_1,\frac{A}{\alpha_1}}(\xi_{\gamma_1}) H_{p_2,\frac{A}{\alpha_1}} (\xi_{\gamma_2}) : p_1+p_2 = 2, \gamma_1<\gamma_2 \} \, .$$

3.   EQUIVALENCE AND SINGULARITY OF SIP's

In this section we combine the representation of SIP's (Theorem 1.4) and
the dichotomy of GP's to the problem of discriminating between two SIP's.
The discrimination problem is completely solved by identifying the Lebesgue
decomposition of the distributions of the two processes and the Radon-Nikodym
derivative of the absolutely continuous part. Two processes are called equiv-
alent ($\sim$), resp. singular ($\perp$), if their induced measures on $\mathcal{B}(\mathbb{R}^T)$ are equiv-
alent ($\sim$), resp. singular ($\perp$). The discrimination problem for two second order
SIP's: $SIP(0,m;F)$ and $SIP(m,S;G)$ is fully resolved by first noting that
without loss of generality we may assume that either $GP(0,R) \sim GP(m,S)$ or else
$GP(0,\alpha R) \perp GP(m,\beta S)$ for all $\alpha,\beta > 0$.   When $GP(0,R) \sim GP(m,S)$, the Lebesgue
decomposition of $SIM(m,S;G)$ with respect to $SIM(0,R;F)$ is given in Theorem
3.1 and the Radon-Nikodym derivative of the absolutely continuous part in
Theorem 3.4. In particular, if $GP(0,R) \sim GP(m,S)$, then $dF \sim dG$ implies
$SIP(0,R;F) \sim SIP(m,S;G)$, and $dF \perp dG$ implies $SIP(0,R;F) \perp SIP(m,S;G)$. Theorem
3.2 shows that if $GP(0,\alpha R) \perp GP(m,\beta S)$ for all $\alpha \beta > 0$ then $SIP(0,R;F) \perp SIP(m,S;G)$.
Finally, if $dF \sim dG$ then $SIP(0,R;F)$ and $SIP(m,R;G)$ are either equivalent or
singular, and necessary and sufficient conditions for equivalence along with
an expression for the Radon-Nikodym derivative are given in Theorem 3.5.

For reasons of clarity in this section we attach the underlying probability
measure to the usual notation for expectation, variance, linear space, etc.

We first state the general theorem concerning the equivalence and the
singularity of two GP's (see for instance Pang (1973)). Fix $\alpha > 0$. Let
$X = (X_t, t \in T)$ be a $GP(0,\alpha R)$ on the probability space $(\Omega, \mathcal{B}(X), P_\alpha)$. Let Q_α
be a second probability on $(\Omega, \mathcal{B}(X))$ under which X is a $GP(m,\alpha S)$. Then either

$P_\alpha \sim Q_\alpha$ or $P_\alpha \perp Q_\alpha$. $P_\alpha \sim Q_\alpha$ if and only if the following conditions are satisfied:

(*)   There exist positive constants K_1 and K_2 such that

$$K_2 \, \text{Var}_{P_\alpha} \xi \leq \text{Var}_{Q_\alpha} \xi \leq K_1 \, \text{Var}_{P_\alpha} \xi$$

for every $\xi \in L(X)$, the linear space of all finite linear combinations $\sum_1^N a_n X_{t_n}$ $t_n \in T$.

(**)   There exists $\xi \in H_{P_\alpha}(X)$ such that $m(t) = E_{P_\alpha}(\xi X_t)$, $t \in T$; i.e. $m \in R(R)$, the reproducing kernel Hilbert space of the covariance R.

(***)   If $B_\alpha : H_{P_\alpha}(X) \to H_{P_\alpha}(X)$ is the positive self-adjoint operator defined by $\text{Cov}_{Q_\alpha}(\xi, \eta) = \text{Cov}_{P_\alpha}(\xi, B_\alpha \eta)$ for all ξ, $\eta \in L(X)$, then $B_\alpha - I$ is Hilbert-Schmidt. Moreover if conditions (*), (**) and (***) hold true and if $\{\lambda_n\}$ and $\{\xi_n\}$ denote the sets of eigenvalues and corresponding normalized eigenvectors of B_α, then

(3.1) $$\frac{dQ_\alpha}{dP} = \exp\{B_\alpha^{-1/2}\xi - \frac{1}{2}\text{Var}_{P_\alpha}(B_\alpha^{-1/2}\xi) - \frac{1}{2}\sum[\xi_n^2(\frac{1}{\lambda_n}-1) + \log\lambda_n]\}$$

From now on $X = (X_t, t \in T)$ will be a second order nondegenerated SIP $(0,R;F)$ on $(\Omega, \mathcal{B}(X), P)$. Recall that R is a covariance function and F is a distribution function on $\mathbb{R}^+$ with finite first moment α_1. P is a mixture of Gaussian measures, $P(E) = \int P_\alpha(E)dF(\alpha)$, and under each P_α, X is a GP$(0,\alpha R)$. Let A be the r.v. associated with X and P as in Section 1.

Now consider a second probability measure Q on $(\Omega, \mathcal{B}(X))$ under which X is a second order nondegenerate SIP $(m,S;G)$. Then $Q(E) = \int Q_\alpha(E)dG(\alpha)$ and under each Q , X is a GP$(m,\alpha S)$. Denote the first moment of G by α_1'. We are interested in the equivalence and mutual singularity of the measures P and Q. Since $P_0 \perp Q_0$ if $m \neq 0$, we shall assume throughout this section that

$F(0+) = 0$. Also we may assume without loss of generality that either $P_1 \sim Q_1$ or $P_\alpha \perp Q_\beta$ for all $\alpha, \beta > 0$, since clearly

(3.2)   a SIP $(m, R; F(\alpha))$, is also a SIP $(m, cR; F(c\alpha))$ for every $c > 0$.

___THEOREM 3.1___ *Let* $P_1 \sim Q_1$ *and let* $dG = dG' + dG''$ *be the Lebesgue decomposition of* dG *with respect to* dF *with* $dG' \ll dF$ *and* $dG'' \perp dF$. *Then*

$$Q(E) = \int Q_\alpha(E)dG'(\alpha) + \int Q_\alpha(E)dG''(\alpha)$$

is the Lebesgue decomposition of Q *with respect to* P. *Hence if* $P_1 \sim Q_1$ *and* $dF \sim dG$ *then* $P \sim Q$; *if* $P_1 \sim Q_1$ *and* $dF \perp dG$ *then* $P \perp Q$.

___Proof:___ Note that $P_1 \sim Q_1$ implies $P_\alpha \sim Q_\alpha$ for all $\alpha > 0$. Let $P(E) = 0$. Then $\int P_\alpha(E)dF(\alpha) = 0$ and $P_\alpha(E) = 0$ a.e. $[dF]$. Since $P_\alpha \sim Q_\alpha$ and $dG' \ll dF$, we have $Q_\alpha(E) = 0$ a.e. $[dG']$, and thus $\int Q_\alpha(E)dG'(\alpha) = 0$. This implies $\int Q_\alpha(\cdot)dF(\alpha) \ll P(\cdot)$.

Since $dG'' \perp dF$, there exists $E \in \mathcal{B}(\mathbb{R})$ such that $\int_E dG''(\alpha) = \int_{E'} dF(\alpha) = 0$. Note that $A = \alpha$ a.e. $[Q_\alpha]$ because $A = \alpha$ a.e. $[P_\alpha]$ and $P_\alpha \sim Q_\alpha$. Thus we have

$$P(A \in E') = \int P_\alpha(A \in E')dF(\alpha) = \int_{E'} dF(\alpha) = 0 ,$$

$$\int Q_\alpha(A \in E)dG''(\alpha) = \int_E dG''(\alpha) = 0 ,$$

which imply $\int Q_\alpha(\cdot)dG''(\alpha) \perp P(\cdot)$.   □

The second assertion of Theorem 3.1 was first stated in Gualtierotti (1974) for P and Q SIM's on a separable Hilbert space.

___THEOREM 3.2___ *If* $P_\alpha \perp Q_\beta$ *for all* $\alpha, \beta > 0$ *then* $P \perp Q$.

___Proof:___ The proof is adapted from Pang (1973).

The following remarks will be used without further comment. For $\xi, \eta \in L(X)$, we have

$$E_{P_\alpha} \xi = E_P \xi = 0 \ , \ \text{Cov}_{P_\alpha}(\xi,\eta) = \frac{\alpha}{\alpha_1} \text{Cov}_P(\xi,\eta) \ ,$$

(3.3)

$$E_{Q_\alpha} \xi = E_Q \xi \ , \ \text{Cov}_{Q_\alpha}(\xi,\eta) = \frac{\alpha}{\alpha'_1} \text{Cov}_Q(\xi,\eta) \ .$$

If P_1 and Q_1 satisfy (*) then for every sequence $\{\xi_n\}$ in $H_{P_1}(X)$ there exist versions of ξ_n's such that $\{\xi_n\}$ is also a sequence in $H_P(X)$, $H_{P_\alpha}(X)$, $H_Q(X)$, $H_{Q_\alpha}(X)$ for all $\alpha > 0$, and satisfies (3.3). (The proof of this is similar to that of Theorem 1.3).

Since $P_1 \perp Q_1$, one of the conditions (*), (**) and (***) must be violated. First suppose that (*) is not satisfied; for instance, suppose that there exists no constant K_1 such that $\text{Var}_{Q_1} \xi \leq K_1 \text{Var}_{P_1} \xi$. Then there exists a sequence $\{\xi_n\}$ in $L(X)$ such that $\text{Var}_{P_1} \xi_n = 1$ and $\text{Var}_{Q_1} \xi > n^2$. Then, as $n \to \infty$, we have

$$P(|\xi_n| > \sqrt{n}) \leq \frac{\alpha_1}{n} \to 0 \ ,$$

$$Q(|\xi_n| > \sqrt{n}) = \int Q_\alpha(|\xi_n| > \sqrt{n}) dG(\alpha)$$
$$\geq \int (1 - \int_{-(\alpha n)^{-1/2}}^{(\alpha n)^{-1/2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx) dG(\alpha)$$
$$\to 1$$

which imply $P \perp Q$.

Next suppose (*) holds but (**) does not hold. Then, for each n, there exists $\xi_n \epsilon L(X)$ such that

$$E_{Q_1} \xi_n > (n \ \text{Var}_{Q_1} \xi)^{1/2} \geq (K_2 n \ \text{Var}_{P_1} \xi_n)^{1/2}$$

(see Pang (1973)). Consequently, as $n \to \infty$

$$P(\xi_n > \frac{1}{2} E_{Q_1} \xi_n) \leq P\{\xi_n > \frac{1}{2}(K_2 n \ \text{Var}_{P_1} \xi_n)^{1/2}\} \leq \frac{4\alpha_1}{K_2 n} \to 0 \ ,$$

$$Q(\xi_n > \frac{1}{2} E_{Q_1} \xi_n) \geq Q(|\xi_n - E_Q \xi_n| < \frac{1}{2} E_Q \xi_n)$$
$$\geq Q(|\xi_n - E_Q \xi_n| < \frac{1}{2}(\frac{n}{\alpha'_1} \ \text{Var}_Q \xi_n)^{1/2})$$
$$\geq 1 - \frac{4\alpha'_1}{n} \to 1 \ ,$$

and thus $P \perp Q$.

Finally, suppose that (*) and (**) are satisfied, but (***) is not. We may assume in this case that $m \equiv 0$. We claim that

(3.4) there exists a sequence $\{\xi_n\}$ in $H_{P_1}(X)$ such that $\text{Cov}_{P_1}(\xi_i, \xi_j)$ $= \delta_{ij}$ and $\text{Cov}_{Q_1}(\xi_i, \xi_j) = \mu_i \delta_{ij}$ where $\sum(1-\mu_n)^2 = \infty$ and $\frac{1}{\mu_n} - \mu_n > 0$ for all n or < 0 for all n.

Given this, consider the events

$$E_m = \{\sum_1^m \xi_n^2(\frac{1}{\mu_n} - 1) \le \frac{1}{2}\sum_1^m(\frac{1}{\mu_n} - \mu_n)\} .$$

We shall deal with the case $\frac{1}{\mu_n} - \mu_n > 0$ only. Note that $K_2 \le \mu_n \le K_1$ by (*). For $\alpha \ge 1$ we have

$$P_\alpha(E_m) = P_1\{\sum_1^m \alpha\xi_n^2(\frac{1}{\mu_n} -1) \le \frac{1}{2}\sum_1^m(\frac{1}{\mu_n} - \mu_n)\}$$

$$\le P_1\{|\sum_1^m (\xi_n-1)(\frac{1}{\mu_n} -1)| \ge \frac{1}{2K_1}\sum_1^m(1-\mu_n)^2\}$$

$$\le 8(\frac{K_1}{K_2})^2 \frac{1}{\sum_1^n(1-\mu_n)^2} \to 0$$

and for $\beta \le 1$

$$Q_\beta(E_m) = Q_1\{\sum_1^m \beta\xi_n^2(\frac{1}{\mu_n} -1) \le \frac{1}{2}\sum_1^m (\frac{1}{\mu_n} - \mu_n)\}$$

$$\ge Q_1\{|\sum_1^m (\xi_n^2-\mu_n)(\frac{1}{\mu_n} -1)| \le \frac{1}{2K_1}\sum_1^m (1-\mu_n^2)\}$$

$$\ge 1 - \frac{8K_1^2}{\sum_1^m(1-\mu_n)^2} \to 1 .$$

Thus if $E = \limsup E_m$, then $P_\alpha(E) = 0$ for $\alpha \ge 1$ and $Q_\beta(E) = 1$ for $\beta \le 1$. This implies by (3.2) that for all $a,b > 0$ there exist $E_{ab} \in \mathcal{B}(X)$ such that

$P_\alpha(E_{ab}) = 0$ for $\alpha \geq a$ and $Q_\beta(E_{ab}) = 1$ for $\beta \leq b$. Hence, as $a \to 0$ and $b \to \infty$, we have

$$P(E_{ab}) = \int_{[0,a]} P_\alpha(E_{ab})dF(\alpha) \leq F(a) \to F(0+) = 0,$$

$$Q(E_{ab}) \geq \int_{[0,b]} Q_\beta(E_{ab})dG(\beta) = G(b) \to 1,$$

and consequently $P \perp Q$.

To complete the proof, we now verify (3.4). Two cases are to be considered. First, suppose that $B_1 - I$ is compact but not Hilbert-Schmidt. Let $\{\lambda_n\}$ and $\{\eta_n\}$ be the ergenvalues and the corresponding normalized ergenvectors of $B_1 - I$. We have $\lambda_n \to 0$, $\sum \lambda_n^2 = \infty$ and $B_1 \eta_n = (1+\lambda_n)\eta_n$. By choosing a suitable subsequence of $\{\eta_n\}$, we obtain the desired sequence $\{\xi_n\}$ in (3.4). Second, suppose that $B_1 - I$ is not compact. Being invertible (by (*)), B_1 is not compact and thus its essential spectrum is not $\{0\}$. Furthermore, since $B_1 - I$ is not compact, there is at least one point $\mu \neq 1$ in the essential spectrum of B_1, and thus also in the essential numerical range of B. (For a nice discussion of essential spectrum and essential numerical range see Fillmore et. al. (1972)). Now by a known result in operator theory (Lemma 2 in Anderson and Stampfli (1971) stated for a separable Hilbert space but true for a nonseparable Hilbert space as well), there exists an orthonormal sequence $\{\eta_n\}$ in $H_{P_1}(X)$ such that $\text{Cov}(\eta_i, B\eta_j) = \mu_i \delta_{ij}$ and $\mu_n \to \mu$. Again $\{\xi_n\}$ in (3.4) is obtained by choosing a suitable subsequence of $\{\eta_n\}$. □

Theorem 3.2 is not true in general, without the Gaussian assumption on P_α and Q_β, since there are uncountably many measures involved. For example, if P_α is the uniform measure on $[0,1]$ for each α, if Q_β is the one point mass at β for each β, and if $dF \sim dG \sim P_1$, then $P_\alpha \perp Q_\beta$ for all α and β, but $P \sim Q$.

Now suppose that $P_1 \sim Q_1$ and let $Q'(\cdot) = \int Q_\alpha(\cdot)dG'(\alpha)$ be the absolutely continuous part of Q given in Theorem 3.1. We will calculate the Radon-Nikodym derivative $\frac{dQ'}{dP}$. Theorems 3.1 and 3.2, together with the expression of $\frac{dQ'}{dP}$, provide a complete solution to the problem of discriminating between two SIP's. To this end, we prepare the following

Lemma 3.3 Let $\rho_\alpha = \frac{dQ_\alpha}{dP_\alpha}$ *a.s.* $[P_\alpha]$. *If* $\rho_{A(\omega)}(\omega)$ *is a measurable function, then*

$$\frac{dQ'}{dP}(\omega) = \rho_{A(\omega)} \frac{dG'}{dF}(A(\omega)) \quad a.s. \ [P].$$

Proof: For every $E \in B(X)$, we have

$$\int_E \rho_A \frac{dG'}{dF}(A)dP = \int \{\int_E \rho_\alpha \frac{dG'}{dF}(\alpha)dP_\alpha\} \ dF(\alpha)$$

$$= \int Q_\alpha(E)dG'(\alpha) = Q'(E)$$

as required.   □

Note that the measurability of $\rho_{A(\omega)}(\omega)$ is not automatic since each ρ_α can be arbitrarily changed on a set of P_α-measure zero.

Theorem 3.4 *Suppose* $P_1 \sim Q_1$. *Then*

(i) *there exists* $\xi \in H_P(X)$ *such that* $m(t) = \frac{1}{\alpha_1} E_P(\xi X_t)$, $t \in T$;

(ii) *there exists a self-adjoint positive operator* B *on* $H_p(X)$ *defined by*

$$\frac{1}{\alpha_1'} Cov_Q(\xi, \eta) = \frac{1}{\alpha_1} Cov_P(\xi, B\eta)$$

for all $\xi, \eta \in L(X)$, *and such that* B - I *is Hilbert-Schmidt.*

Moreover if $\{\lambda_n\}$ *and* $\{\xi_n\}$ *is the set of eigenvalues of* B *and their corresponding eigenvectors with norms* $\alpha_1^{1/2}$, *then*

$$(3.5) \quad \frac{dQ'}{dP} = \exp \{\frac{1}{A} B^{-1/2}\xi - \frac{1}{2\alpha_1 A} Var_P(B^{-1/2}\xi)$$

$$- \frac{1}{2} \sum_1^\infty [\frac{\xi_n^2}{A}(\frac{1}{\lambda_n}-1)+\log\lambda_n]\} \cdot \frac{dG'}{dF}(A) .$$

Proof: Let $X = A^{1/2}Y$ be the canonical representation of X, where Y is a $GP(0,R)$ under P and under each P_α $(\alpha > 0)$. It is clear that each $\xi \epsilon H_p(X)$ or $H_{p_\alpha}(X)$ has a version of the form $A^{1/2}\eta$ with $\eta \epsilon H_p(Y)$. In this sense, we have $H_p(X) = A^{1/2}H(Y) = H_{p_\alpha}(X)$. Therefore every operator on $H_p(X)$ induces an operator on $H_{p_\alpha}(X)$ in an obvious way, and vice versa.

Since $P_1 \sim Q_1$, there exists a $\xi \epsilon H_{p_1}(x)$ such that $m(t) = E_{p_1}(\xi X_t), t \epsilon T$, and a self-adjoint positive operator B_1 on $H_{p_1}(X)$ defined by $Cov_{Q_1}(\xi,\eta) = Cov_{p_1}(\xi, B_1\eta), \xi, \eta \epsilon L(X)$. A simple computation shows that ξ satisfies (i) and that B, the operator on $H_p(X)$ induced by B_1, satisfies (ii).

By Lemma 3.3, (3.5) will follow if we show

$$(3.6) \quad \frac{dQ_\alpha}{dP_\alpha} = \exp\left\{\frac{1}{\alpha} B^{-1/2}\xi - \frac{1}{2\alpha_1\alpha} Var_P (B^{-1/2}\xi) - \frac{1}{2} \sum_1^\infty \left[\frac{\xi_n^2}{\alpha}\left(\frac{1}{\lambda_n}-1\right)+\log\lambda_n\right]\right\} \text{ a.s. } [P_\alpha].$$

Let B_α be the operator on $H_{p_\alpha}(X)$ induced by B. Again, a simple computation shows that $\alpha^{-1}\xi$ satisfies condition (**), and that B_α satisfies condition (***) and has eigenvalues $\{\lambda_n\}$ and corresponding eigenvectors $\{\alpha^{-1/2}\xi_n\}$. Thus we have by (3.1)

$$\frac{dQ_\alpha}{dP_\alpha} = \exp\left\{\frac{1}{\alpha} B_\alpha^{-1/2}\xi - \frac{1}{2\alpha^2} Var_{P_\alpha} (B_\alpha^{-1/2}\xi) - \frac{1}{2} \sum \left[\frac{\xi_n^2}{\alpha}\left(\frac{1}{\lambda_n}-1\right)+\log\lambda_n\right]\right\} \text{ a.s.} [P_\alpha]$$

which is equivalent to (3.6) since $B_\alpha^{-1/2}\xi = B^{-1/2}\xi$ a.s. $[P_\alpha]$. ☐

Theorem 3.5   _Suppose R = S and dF ~ dG.   Then either P ~ Q or_

P ⊥ Q, and P ~ Q if and only if

(3.7) $m(t) = E_p(\xi X_t), t \in T,$

for some $\xi \in H_p(X)$. _In the case of equivalence,_

(3.8) $\frac{dQ}{dP} = \exp \{\frac{\alpha_1}{A} \ (\xi - \frac{1}{2} \ Var_p \xi) \ . \ \frac{dG}{dF} \ (A)$

<u>Proof</u>: Clearly if (3.7) holds we have $P_1 \sim Q_1$, and if (3.7) does not hold

we have $P_\alpha \perp Q_\beta$ for all α, β. The first assertion follows then from

Theorems 3.1 and 3.2. (3.8) follows from Theorem 3.4 by comparing (3.7)

with (3.6) and noting that R = S implies B = I.   □

Sytaya (1969) derived (3.8) for P and Q SIM's on a separable Hilbert

space and F = G.

4. NONLINEAR ESTIMATION AND PREDICTION

Using the tensor product structure of the nonlinear space we solve the general nonlinear estimation problem for SIP's, and in particular for GP's, in the sense that we reduce the nonlinear problem to a standard linear estimation problem, the theory of which is well developed. Also we derive a lower bound for the mean square error of the nonlinear prediction for a certain class of prediction problems.

4.1 Nonlinear Estimation

Let $X = (X_t, t \epsilon T)$ be a second order process with zero mean. Consider the following estimation problem: we observe X_t for $t \epsilon S$, a subset of T, and we want to estimate an L_2-functional θ of X based on the observations. We are interested in finding the best estimate $\hat{\theta}$, an L_2-functional of $(X_t, t \epsilon S)$ which minimizes the mean square error of estimation $E(\hat{\theta} - \theta)^2$, and it is well known that

$$\hat{\theta} = E(\theta | X_t, t \epsilon S) .$$

In general, $\hat{\theta}$ is extremely difficult to determine. However, if X is a SIP, we have a complete solution. In formulating the main result we will use the notation of Section 2, and identify $L_2(X)$ with $L_2(A) \otimes (\underset{p \geq 0}{\oplus} H^{\widetilde{\otimes}p}(X))$ by Theorem 2.4. We let $L_2(X;S) = L_2(X_t, t \epsilon S)$ and $H(X;S) = H(X_t, t \epsilon S)$.

<u>THEOREM 4.1</u> *Let* X *be a nondegenerate SIP* $(0,R;F)$ *with* $F(0+) = 0$ *and let* $\theta \epsilon L_2(X)$ *have the following orthogonal development*

$$\theta = \sum_{\substack{n \geq 1 \\ p \geq 0}} \sum_{\substack{p_1 + \ldots + p_k = p \\ \gamma_1 < \ldots < \gamma_k}} {}_n a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k} e_n(A) \otimes (\xi_{\gamma_1}^{\widetilde{\otimes}} \widetilde{\otimes} \ldots \widetilde{\otimes} \xi_{\gamma_k}^{\widetilde{\otimes}p_k}) .$$

Suppose that $(X_t, \ t\epsilon S)$ *is nondegenerate. Then*

$$\hat{\theta} = \sum_{\substack{n\geq 1 \ p_1 + \ldots + p_k = p \\ p\geq 0 \ \gamma_1 < \ldots < \gamma_k}} {}_n a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k} e_n(A) \otimes (\hat{\xi}_{\gamma_1}^{\widetilde{\otimes}p_1} \widetilde{\otimes} \ldots \widetilde{\otimes} \hat{\xi}_{\gamma_k}^{\widetilde{\otimes}p_k}) \ .$$

where for $\xi\epsilon H(X)$,

$$\hat{\xi} = Proj_{H(X;S)}\xi \ (=E(\xi|X_t, \ t\epsilon S)) \ .$$

<u>Proof</u>: Recall the definition of the r.v. A and observe that it is independent of the choice of the defining sequence $\{\xi_i\}$. It then follows that $A\epsilon L_2(X;S)$ and every $\rho\epsilon L_2(X;S)$ has the orthogonal development

$$\rho = \sum_{\substack{q\geq 0 \ q_1 + \ldots + q_j = q \\ m\geq 1 \ \delta_1 < \ldots < \delta_j}} {}_m b_{\delta_1 \ldots \delta_j}^{q_1 \ldots q_j} e_m(A) \otimes (\eta_{\delta_1}^{\widetilde{\otimes}q_1} \widetilde{\otimes} \ldots \widetilde{\otimes} \eta_{\delta_j}^{\widetilde{\otimes}q_j}) \ .$$

where $\{\eta_\delta, \ \delta\epsilon\Delta\}$ (Δ linearly ordered) is a CONS in $H(X;S)$.

We have

$$\hat{\theta} = E(\theta|X_t, \ t\epsilon S) = Proj_{L_2(X;S)}\theta$$

$$= \sum_{\substack{p\geq 0 \ p_1 + \ldots + p_k = p \\ n\geq 0 \ \gamma_1 < \ldots < \gamma_k}} {}_n a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k} Proj_{L_2(X;S)} e_n(A) \otimes (\xi_{\gamma_1}^{\widetilde{\otimes}p_1} \widetilde{\otimes} \ldots \widetilde{\otimes} \xi_{\gamma_k}^{\widetilde{\otimes}p_k}) \ .$$

Thus to show the theorem it suffices to show that

$$Proj_{L_2(X;S)} e_n(A) \otimes (\xi_{\gamma_1}^{\widetilde{\otimes}p_1} \widetilde{\otimes} \ldots \widetilde{\otimes} \xi_{\gamma_k}^{\widetilde{\otimes}p_k}) = e_n(A) \otimes (\hat{\xi}_{\gamma_1}^{\widetilde{\otimes}p_1} \widetilde{\otimes} \ldots \widetilde{\otimes} \hat{\xi}_{\gamma_k}^{\widetilde{\otimes}p_k}) \ .$$

If $\xi_1, \ldots, \xi_p \epsilon H(X)$ and $\eta_1, \ldots, \eta_p \epsilon H(X;S)$ then it follows from $\langle \xi_i, \eta_j \rangle_{H(X)} = \langle \hat{\xi}_i, \eta_j \rangle_{H(X)}$ that

$$\langle \xi_1 \widetilde{\otimes} \ldots \widetilde{\otimes} \xi_p, \eta_1 \widetilde{\otimes} \ldots \widetilde{\otimes} \eta_p \rangle_{H^{\widetilde{\otimes}p}(X)} = \langle \hat{\xi}_1 \widetilde{\otimes} \ldots \widetilde{\otimes} \hat{\xi}_p, \eta_1 \widetilde{\otimes} \ldots \widetilde{\otimes} \eta_p \rangle_{H^{\widetilde{\otimes}p}(X)}$$

and hence for each $\rho\epsilon L_2(X;S)$ we have

$$
\langle e_n(A) \; \tilde{\otimes} \; (\xi_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \xi_{\gamma_k}^{\tilde{\otimes}p_k}), \; p \rangle_{L_2(X)}
$$

$$
= \sum_m b \begin{matrix} q_1 \ldots q_j \\ \delta_1 \ldots \delta_j \end{matrix} \langle e_n(A) \otimes (\xi_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \xi_{\gamma_k}^{\tilde{\otimes}p_k}), \; e_m(A) \otimes (\eta_{\delta_1}^{\tilde{\otimes}q_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \eta_{\delta_j}^{\tilde{\otimes}q_j}) \rangle_{L_2(X)}
$$

$$
= \sum_m b \begin{matrix} q_1 \ldots q_j \\ \delta_1 \ldots \delta_j \end{matrix} \langle e_n(A) \otimes (\hat{\xi}_{\gamma_1}^{\tilde{\otimes}p_1} \tilde{\otimes} \ldots \tilde{\otimes} \; \xi_{\gamma_k}^{\tilde{\otimes}p_k}), \; e_m(A) \otimes (\eta_{\delta_1}^{\tilde{\otimes}q_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \eta_{\delta_j}^{\tilde{\otimes}q_j}) \rangle_{L_2(X)}
$$

$$
= \langle e_n(A) \otimes (\hat{\xi}_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \hat{\xi}_{\gamma_k}^{\tilde{\otimes}p_k}), \; \rho \rangle_{L_2(X)} \; .
$$

Since $e_n(A) \otimes (\hat{\xi}_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \hat{\xi}_{\gamma_k}^{\tilde{\otimes}p_k}) \in L_2(X;S)$, the result follows. $\qquad\square$

A similar result can be obtained when $F(0+) > 0$. Of particular interest is the case where X is in fact a Gaussian process. In this case we have $A = 1 = \alpha_1$, $L_2(X) \cong \bigoplus_{p \geq 0} H^{\tilde{\otimes}p}(X)$, $\bar{Q}_p \cong H^{\tilde{\otimes}p}(X)$; and Theorem 4.1 yields the following.

<u>COROLLARY 4.2</u>   Let X *be a* $GP(0,R)$ *and let* $\theta \in L_2(X)$ *have the following orthogonal development*

$$
\theta = \sum_{p \geq 0} \sum_{\substack{p_1 + \ldots + p_k = p \\ \gamma_1 < \ldots < \gamma_k}} a \begin{matrix} p_1 \ldots p_k \\ \gamma_1 \ldots \gamma_k \end{matrix} \xi_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \xi_{\gamma_k}^{\tilde{\otimes}p_k} \; .
$$

Then

$$
\hat{\theta} = \sum_{p \geq 0} \sum_{\substack{p_1 + \ldots + p_k = p \\ \gamma_1 < \ldots < \gamma_k}} a \begin{matrix} p_1 \ldots p_k \\ \gamma_1 \ldots \gamma_k \end{matrix} \hat{\xi}_{\gamma_1}^{\tilde{\otimes}p_1} \; \tilde{\otimes} \ldots \tilde{\otimes} \; \hat{\xi}_{\gamma_k}^{\tilde{\otimes}p_k}
$$

where for $\xi \in H(X)$,

$$
\hat{\xi} = \text{Proj}_{H(X;S)} \xi \quad (= E(\xi | X_t, \; t \in S)) \; .
$$

Consequently,

$$
\text{Proj}_{\bar{Q}_p} \cdot \text{Proj}_{L_2(X;S)} = \text{Proj}_{L_2(X;S)} \cdot \text{Proj}_{\bar{Q}_p} = \text{Proj}_{\bar{Q}_p(S)}
$$

where $\bar{Q}_p(S)$ denotes the p-th homogeneous chaos of $(X_t, t\epsilon S)$.

COROLLARY 4.3   If X is a nondegenerate SIP $(0,R;F)$ and if $\int e^{t\alpha}dF(\alpha) < \infty$ for all $t \epsilon \mathbb{R}$, then for all $\xi\epsilon H(X)$,

$$E(H_{p,\frac{A}{\alpha_1}E\xi^2}(\xi)|X_t, t\epsilon S) = H_{p,\frac{A}{\alpha_1}E\hat{\xi}^2}(\hat{\xi}) ,$$

$$E(\exp\{\xi - \frac{A}{2\alpha_1}E\xi^2\}|X_t, t\epsilon S) = \exp\{\hat{\xi} - \frac{A}{2\alpha_1}E\hat{\xi}^2\} .$$

In particular, if X is a GP$(0,R)$ and $\xi\epsilon H(X)$, then

$$E(H_{p,E\xi^2}(\xi)|X_t, t\epsilon S) = H_{p,E\hat{\xi}^2}(\hat{\xi}) ,$$

$$E(\exp\{\xi - \tfrac{1}{2}E\xi^2\}|X_t, t\epsilon S) = \exp\{\hat{\xi} - \tfrac{1}{2}E\hat{\xi}^2\} .$$

If X is a zero mean Gaussian martingale then

$$H_{p,EX_t^2}(X_t), \quad \exp\{X_t - \tfrac{1}{2}EX_t^2\}$$

are martingales.

Proof: Only the first assertion requires proof and we shall prove it for the case $F(0+) = 0$ only. $\int e^{t\alpha}dF(\alpha) < \infty$, $t \epsilon \mathbb{R}$, implies $A^{p/2}\epsilon L_2(A)$. Thus

$$\theta_1 = H_{p,\frac{A}{\alpha_1}E\xi^2}(\xi) = (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{p}{2}} \cdot (\frac{1}{p!})^{\frac{1}{2}}(\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p,\frac{A}{\alpha_1}E\xi^2}(\xi) \qquad \text{(by (2.4))}$$

$$= (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{p}{2}} \otimes \xi^{\tilde{\otimes}p} \epsilon L_2(A) \otimes H^{\tilde{\otimes}p}(X) \subset L_2(X)$$

and by Theorem 4.1

$$\hat{\theta}_1 = (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{p}{2}} \otimes \hat{\xi}^{\tilde{\otimes}p} = H_{p,\frac{A}{\alpha_1}E\hat{\xi}^2}(\hat{\xi}) .$$

Now write

$$\theta_2 = \exp\{\xi - \frac{A}{2\alpha_1} E\xi^2\} = \sum_{p \geq 0} \frac{1}{p!} H_{p, \frac{A}{\alpha_1}E\xi^2}(\xi) \quad,$$

the convergence of the series being pointwise. It follows from the expression of θ_1 that all terms in the above sum are mutually orthogonal and each term has square norm $\frac{\alpha_p}{\alpha_1^p} \cdot \frac{(E\xi^2)^p}{p!}$, where $\alpha_p = \int \alpha^p dF(\alpha)$. Therefore the series converges also in $L_2(X)$ and

$$\hat{\theta}_2 = \sum_{p \geq 0} \frac{1}{p!} E(H_{p, \frac{A}{\alpha_1}E\xi^2}(\xi)|X_t, t \in S) = \sum_{p \geq 0} \frac{1}{p!} H_{p, \frac{A}{\alpha_1}E\hat{\xi}^2}(\hat{\xi})$$

$$= \exp\{\hat{\xi} - \frac{A}{2\alpha_1} E\hat{\xi}^2\} \quad. \qquad \qquad \square$$

For X a Wiener process it is well known that $(X_t^2 - t, t \geq 0)$ and $(\exp\{X_t - \frac{t}{2}\}, t \geq 0)$ are martingales.

If X is a zero mean Gaussian process and $T = (-\infty, \infty)$ (or any interval) then by Corollary 4.3 we have that for all $s \leq t$,

$$E(H_{p, EX_t^2}(X_t)|X_u, u \leq s) = H_{p, E\hat{X}_{t,s}^2}(\hat{X}_{t,s})$$

where

$$\hat{X}_{t,s} = E(X_t|X_u, u \leq s) \quad.$$

An expression for $\hat{X}_{t,s}$ can always be obtained via the Cramér-Hida representation of X:

$$X_t = \sum_{n=1}^{N} \int_{-\infty}^{t} f^{(n)}(t,u) dZ_u^{(n)} \quad.$$

Then we have

$$\hat{X}_{t,s} = \sum_{n=1}^{N} \int_{-\infty}^{s} f^{(n)}(t,u) dZ_u^{(n)} \quad.$$

The case with $p = 2$, i.e. the L_2-functional $X_t^2 - EX_t^2$, is considered in Hida and Kallianpur (1975) for a special class of Gaussian processes X. It should be clear that whenever a simple expression is available for $\hat{X}_{t,s}$, then a simple expression is also found for the nonlinear predictor of $H_{p,EX_t^2}(X_t)$.

We close this section with a simple example. Let $(X_t, -\infty < t < \infty)$ be a stationary Gaussian process with $EX_t = 0$, $EX_t^2 = 1$, and continuous covariance $R(t)$, which is reciprocal or quasi-Markov on $[0,T]$. Then it is known [13, 5] that $R(t)$, $0 \le t \le T$, has one of the following three forms: $Ae^{-at} + (1-A)e^{at}$ where $a > 0$, $A > \frac{1}{2}$, $T \le (2a)^{-1}\log(A/|A-1|)$; $\cos at$ where $a > 0$, $T \le \pi/a$; $1-at$ where $0 \le a \le 2/T$. We want to estimate θ, an L_2-functional of X_t, $0 \le u < t < v \le T$, based on observations X_s, $s \epsilon S = [0,u] \cup [v,T]$. Since X is reciprocal or quasi-Markov on $[0,T]$ we have

$$\hat{X}_t = E(X_t | X_s, s \epsilon S) = \alpha X_u + \beta X_v$$

where

$$\alpha = \frac{R(u-t)-R(v-t)R(u-v)}{1-R^2(u-v)} \quad , \qquad \beta = \frac{R(v-t)-R(u-t)R(u-v)}{1-R^2(u-v)} \quad .$$

Since θ is an L_2-functional of X_t, it has the orthogonal development $\theta = \sum a_p H_p(X_t)$, and thus the best estimate $\hat{\theta}$ of θ is given by

$$\hat{\theta} = \sum_{p \ge 0} a_p H_{p,EX_t^2}(\hat{X}_t) = \sum_{p \ge 0} a_p H_{p,\alpha^2+\beta^2+2\alpha\beta R(u-v)}(\alpha X_u + \beta X_v) \quad .$$

4.2 Nonlinear Prediction

Consider the following prediction problem: Let $X = (X_t, t \epsilon T)$, T an interval, be a second order process and let $Y = (Y_t = \theta_t(X_t), t \epsilon T)$ with θ_t a real function

such that $EY_t = 0$ and $EY_t^2 < \infty$ for all $t \epsilon T$. Suppose that on the basis of the past values of Y up to time t we want to find the best prediction of the future value $Y_{t+\tau}$, $\tau > 0$. Two predictors are of special interest: the optimal linear predictor $\hat{Y}_\ell(t,\tau)$ and the nonlinear predictor $\hat{Y}_{n\ell}(t,\tau)$. The optimality is in the sense of minimizing the mean square error within the class of all linear and nonlinear predictors respectively. It is well known that

$$\hat{Y}_{n\ell}(t,\tau) = E(Y_{t+\tau}|Y_s, \ s\leq t) \ , \qquad \hat{Y}_\ell(t,\tau) = \text{Proj}_{H(Y_s, s\leq t)} Y_{t+\tau} \ .$$

The corresponding mean square predictor errors are denoted by

$$\sigma_{n\ell}^2(t,\tau) = E[Y_{t+\tau}-\hat{Y}_{n\ell}(t,\tau)]^2 \ , \quad \sigma_\ell^2(t,\tau) = E[Y_{t+\tau}-\hat{Y}_\ell(t,\tau)]^2 \ .$$

Now introduce a "super predictor" $\hat{Y}_s(t,\tau)$ as the nonlinear predictor of $Y_{t+\tau}$ based on $(X_s, \ s\leq t)$, i.e.

$$\hat{Y}_s(t,\tau) = E(Y_{t+\tau}|X_s, \ s\leq t) \ ,$$

and denote its mean square error by $\sigma_s^2(t,\tau)$. It is clear that

(4.1) $$\sigma_s^2(t,\tau) \ \leq \ \sigma_{n\ell}^2(t,\tau) \ \leq \ \sigma_\ell^2(t,\tau)$$

and thus σ_s^2 provides a lower bound for the mean square errors of linear and nonlinear prediction. If X is a SIP then σ_s^2 can be obtained as in Section 4.1 by solving an estimation problem. If, in addition, θ_t is one-to-one for each t then the σ-fields generated by X_t and Y_t coincide. In this case $\hat{Y}_{n\ell}(t,\tau) = \hat{Y}_s(t,\tau) = E(Y_{t+\tau}|X_s, \ s\leq t)$, and the nonlinear predictor can be again obtained by solving an estimation problem.

In the important case where $X = (X_t, \ t \ \epsilon \ \mathbb{R})$ is a zero mean stationary Gaussian process with covariance function $R(t,s) = R(t-s)$ and $\theta_t = \theta$ we can calculate the lower bound $\sigma_s^2(t,\tau) = \sigma_s^2(\tau)$ as follows. Write

$$(4.2) \qquad Y_t = \theta(X_t) = \sum_{p=1}^{\infty} a_p H_{p,\sigma^2}(X_t)$$

where $\sigma^2 = EX_t^2$. Note that for all $\xi, \eta \in H(X)$,

$$EH_{p,E\xi^2}(\xi)H_{p,E\eta^2}(\eta) = p! < \xi^{\tilde{\otimes}p}, \eta^{\tilde{\otimes}p} > = p!\{E(\xi\eta)\}^p$$

$$EH_{p,E\xi^2}(\xi)H_{q,E\eta^2}(\eta) = 0 \qquad \text{if } p \neq q.$$

Thus we have

$$(4.3) \qquad EY_t Y_s = \sum_{p \geq 1} p! a_p^2 R^p(t-s) .$$

Let $\hat{X}(t,\tau) = E(X_{t+\tau}|X_s, s \leq t)$ be the optimal nonlinear predictor of $X_{t+\tau}$ (which is also the optimal linear predictor since X is Gaussian), and $\sigma_0^2(\tau)$ be the mean square error. Then by Corollary 4.3,

$$(4.4) \qquad \hat{Y}_s(t,\tau) = \sum_{p \geq 1} a_p H_{p,E\hat{X}^2(t,\tau)}(\hat{X}(t,\tau))$$

and hence

$$\sigma_s^2(\tau) = E(Y_{t+\tau} - \hat{Y}_s(t,\tau))^2 = EY_{t+\tau}^2 - E(\hat{Y}_s(t,\tau))^2$$

$$(4.5) \qquad = \sum_{p \geq 1} p! a_p^2 \sigma^{2p} - \sum_{p \geq 1} p! a_p^2 [\sigma^2 - \sigma_0^2(\tau)]^p$$

$$= \sum_{p \geq 1} p! a_p^2 \{\sigma^{2p} - [\sigma^2 - \sigma_0^2(\tau)]^p\} .$$

It is well known from the general theory of stationary processes (e.g. Doob (1953), Rozanov (1967)] that $\sigma_0^2(\tau)$ can be obtained analytically (if not explicitly) through the Wiener-Paley factorization theorem if X is linearly regular, i.e. $\cap_t H(X_t, s \leq t) = \{0\}$. When X is mean square continuous and linearly regular we now show that so is Y, and hence $\sigma_\ell^2(\tau)$ can be obtained analytically. Y is clearly stationary and its mean square continuity follows from the continuity of R and (4.3). The linear regularity of Y follows from the fact that

a Gaussian process is linearly regular if and only if its remote past is trivial. Here we give a purely geometric proof of this property; for a proof using Kolmogorov's zero-one law see Rosanov (1967) and Ibragimov and Linnik (1971).

THEOREM 4.4 Let $X = (X_t, t \in \mathbb{R})$ *be a zero mean Gaussian process ,* $B_t = B(X_s, s \leq t)$, $t \in \mathbb{R}$, *and* $B_{-\infty} = \cap_t B_t$. *Then X is linearly regular if and only if* $B_{-\infty}$ *is trivial.*

<u>Proof</u>: The "if" part is clear. For the "only if" part, note first that the triviality of $B_{-\infty}$ is equivalent to the condition

$$\cap_t L_2(X_s, s \leq t) = \{0\} .$$

For simplicity we write $L_t = L_2(X_s, s \leq t)$ and $H_t = H(X_s, s \leq t)$. Thus we need to show that $H_{-\infty} = \cap_t H_t = \{0\}$ implies $L_{-\infty} = \cap_t L_t = \{0\}$. So assume that $H_{-\infty} = \{0\}$ and let $\theta \in L_{-\infty}$. We will show that $\theta = 0$.

Fix a sequence $\{t_n, n=1,2,\ldots\}$ decreasing to $-\infty$. The family of subspaces H_{t_n} has the property that $H_{t_m} \subseteq H_{t_n}$ for $m < n$ and $\cap_n H_{t_n} = \{0\}$. Thus it follows readily (see e.g. Rozanov (1967), pp. 53, 56) that $\lim_n \text{Proj}_{H_{t_n}} \xi = 0$ for all $\xi \in H(X)$, and hence

$$H_{t_1} = \overset{\infty}{\underset{n=1}{\oplus}} D_n$$

where D_n is the orthogonal complement of $H_{t_{n+1}}$ in H_{t_n}, i.e., $H_{t_n} = H_{t_{n+1}} \oplus D_n$. In each D_n pick a CONS; then pool all these CONS's together to get a CONS $\{\xi_\gamma, \gamma \in \Gamma\}$ in H_{t_1}. Observe that

(4.6) $$\text{Proj}_{H_{t_n}} \xi_\gamma = 0 \text{ or } \xi_\gamma .$$

Since $\theta \in L_{-\infty} \subset L_{t_1}$, it has the orthogonal development

$$\theta = \sum_{\substack{p \geq 0}} \sum_{\substack{p_1 + \ldots + p_k = p \\ \gamma_1 < \ldots < \gamma_k}} a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k} \xi_{\gamma_1}^{\tilde{\otimes}p_1} \tilde{\otimes} \ldots \tilde{\otimes} \xi_{\gamma_k}^{\tilde{\otimes}p_k} \quad ,$$

and by Corollary 4.2,

$$\theta = \text{Proj}_{L_{t_n}} \theta = \sum \sum a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k} (\text{Proj}_{H_{t_n}} \xi_{\gamma_1})^{\tilde{\otimes}p_1} \tilde{\otimes} \ldots \tilde{\otimes} (\text{Proj}_{H_{t_n}} \xi_{\gamma_k})^{\tilde{\otimes}p_k}$$

which together with (4.6) yields

$$(4.7) \qquad E\theta^2 = \sum \sum (a_{\gamma_1 \ldots \gamma_k}^{p_1 \ldots p_k})^2 || (\text{Proj}_{H_{t_n}} \xi_{\gamma_1})^{\tilde{\otimes}p_1} \tilde{\otimes} \ldots \tilde{\otimes} (\text{Proj}_{H_{t_n}} \xi_{\gamma_k})^{\tilde{\otimes}p_k} ||^2 .$$

Now let $n \to \infty$ in (4.7). The limit can be taken inside the summation since each summand is bounded by $|| \xi_{\gamma_1}^{\tilde{\otimes}p_1} \tilde{\otimes} \ldots \tilde{\otimes} \xi_{\gamma_k}^{\tilde{\otimes}p_k} ||^2$; and the limit of each summand is zero because of (4.6). Consequently $E\theta^2 = 0$, i.e. $\theta = 0$, and the theorem is proved. ☐

Jaglom (1970) has considered the problem of comparing the performance of optimal linear and nonlinear predictors for polynomial functions of certain stationary Markov processes. Donelson and Maltz (1972) studied this problem in detail for polynomial functions of the Ornstein-Uhlenbeck process. The inequality (4.1) plays a central role in such studies. As an example, let X be the Ornstein-Uhlenbeck process, i.e. a zero mean Gaussian process with covariance function $R(t,s) = e^{-|t-s|}$, and let Y be given by (4.2). By the Markov property of X we have

$$\hat{X}(t,\tau) = E(X_{t+\tau} | X_s, \; s \leq t) = e^{-\tau} X_t .$$

Thus it follows from (4.4) and (4.5) that

$$\hat{Y}_s(t,\tau) = \sum_{p \geq 1} a_p H_{p,e^{-2\tau}} (e^{-\tau} X_t) = \sum_{p \geq 1} a_p e^{-p\tau} H_p(X_t) \quad ,$$

$$\sigma_s^2(\tau) = \sum_{p \geq 1} p! a_p^2 (1 - e^{2p\tau})^2 \ .$$

This result, with Y_t a polynomial function of X_t, has been obtained by Donelson and Maltz using a different approach; they also compared σ_s^2 with σ_ℓ^2 and found that these two errors are frequently close to one another. Finally, we remark that if $Y_t = H_p(X_t)$ then

$$\hat{Y}_{n\ell}(t,\tau) = \hat{Y}_s(t,\tau) = e^{-p\tau} Y_t \ ,$$

$$\sigma_{n\ell}^2(\tau) = \sigma_s^2(\tau) = 1 - e^{-2p\tau} \ .$$

REFERENCES

[1] ANDERSON, J.H. AND STAMPFLI, J.G. (1971). Commutators and Compressions. Israel J. Math. 10 433-441.

[2] BESSON, J.L. (1974). Continuité des fonctions aléatoires spheriquement invariantes. Publication du Département de Mathématiques, Lyon, 59-69.

[3] BRETAGNOLLE, J., CASTELLE, D.D. AND KRIVINE, J.L. (1966). Lois stables et espaces L^p. Ann. Inst. H. Poincaré, Ser. B, 2 231-259.

[4] CAMERON, R.H. AND MARTIN, W.T. (1947). The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals. Ann. Math. 48 385-392.

[5] CHAY, S.C. (1972). On quasi-Markov random fields. J. Multivariate Anal. 2 14-76.

[6] DOOB, J.L. (1953). Stochastic Processes. Wiley, New York.

[7] DONELSON, J. III AND MALTZ, F. (1972). A comparison of linear versus non-linear prediction for polynomial functions of the Ornstein-Uhlenbeck process. J. Appl. Probability 9 725-744.

[8] FILLMORE, P.A., STAMPFLI, J.G. AND WILLIAMS, J.P. (1972). On the essential numerical range, the essential spectrum, and a problem of Halmos. Acta Sci. Math. (Szeged) 33 179-192.

[9] GUALTIEROTTI, A.F. (1974). Some remarks on spherically invariant distributions. J. Multivariate Anal. 4 347-349.

[10] HIDA, T. AND KALLIANPUR, G. (1975). The square of a Guassian Markov process and nonlinear prediction. J. Multivariate Anal. 5 451-461.

[11] IBRAGIMOV, I.A. AND LINNIK, Yu.V. (1971). Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen, The Netherlands.

[12] JAGLOM, A.M. (1970). Examples of optimal nonlinear extrapolation of stationary processes. Selected Transl. in Math. Statist. and Probability, 9, American Math. Soc. 273-298.

[13] JAMISON, B. (1970). Reciprocal processes: the stationary Gaussian case. Ann. Math. Statist. 41 1624-1630.

[14] NAGORNYI, V.N. (1974). Random processes for which optimal prediction is linear. Theor. Probability Math. Statist. 1 147-154.

[15] NEVEU, J. (1968). Processus Aléatoires Gaussiens. Les Presses de L'Université de Montréal, Montréal.

[16] PAN, Y-M. (1973). Simple proofs of equivalence conditions for measures induced by Gaussian processes. Selected Transl. in Math. Statist. Probability, 12, Amer. Math. Soc. 109-118.

[17] ROZANOV, Yu.A. (1971). Stationary Random Processes. Holden-Day,
      San Francisco, CA.

[18] SYTAYA, G.N. (1969). On admissible displacements of suspended Gaussian
      measures. Theor. Probability Appl. 14 506-509.

[19] VERSHIK, A.M. (1964). Some characteristic properties of Gaussian stochastic
      processes. Theor. Probability Appl. 9 353-356.